

Poverty persistence among Belgian elderly in the transition from work to retirement : an empirical analysis

M. Maes

Discussion Paper 2008-42

Département des Sciences Économiques
de l'Université catholique de Louvain



UCL

Poverty persistence among Belgian elderly in the transition from work to retirement: an empirical analysis

Marjan Maes

UCLouvain HUBrussel

Abstract

On the basis of a longitudinal administrative dataset(1991-2002) merged with the Census of 2001 and the National Register, the majority of the poor elderly in Belgium appear to be persistently poor. The question arises why this might be so. To the extent that individual characteristics such as low abilities persist over time, they may also be the reason that individuals persist in poverty over time. In that case, one expects that once individual characteristics are controlled for, duration dependence in poverty becomes spurious. The alternative possibility is that poverty experience has a causal impact on future poverty. This may be because of a poverty trap: people may be given an incentive not to work while at the same time they slip into poverty. Or this may be due to depreciation of human capital or loss of motivation. The reasons for dependence in poverty are of interest for developing effective poverty reducing measures since true dependence would suggest to focus on stigma and adverse work incentives while spurious dependence would suggest to change individual's characteristics. The simultaneous estimation of a multiple-spell discrete-time hazard model of transitions in and out of poverty, that allows for unobserved effects and a significant initial condition problem, lends strong empirical support for true duration dependence in poverty.

This suggestion sounds reasonable since in Belgium elderly unemployed are exempted from the search for a job and thus easily exposed to depreciation of human capital and employers are reluctant to invest in the human capital of older workers. In addition in Belgium both employers and the government design retirement pathways that give elderly strong incentives to leave the labour market as soon as possible.

Keywords: poverty dynamics; poverty persistence; early retirement; work disincentives; multiple spell discrete-time hazard model

JEL codes: J14; J26; C41; I32

1. Introduction

OECD studies¹ and other studies² argued several times that working is an effective means of staying out of poverty. At a first glance this seems difficult to reconcile with the fact that most OECD countries designed social security systems that strongly encourage

¹ OECD(2005a), OECD(2005b), OECD Jobs Strategy(1994), OECD(1997), Förster(2000), Casey-Yamada(2002).

² On the basis of EU-SILC(2004-2005), Zaidi et al.(2006): "We also find that a large proportion of elderly have a high risk of persistent poverty. This can be true by default, since the elderly have little opportunities to enhance their income position in post-retirement life. Thus, the most effective policy intervention to enhance incomes of the elderly will be to increase incentives to work." See also Bardasi-Jenkins-Rigg(2000) for the UK.

elderly to stop working as soon as possible³. Given the fact that youth unemployment is left unaffected by early retirement of older workers⁴ and that early retirement threatens the financial sustainability of social security systems, it becomes hard to understand why the Belgian government continues to encourage early retirement, especially if the latter would increase the risk of poverty among those who retire early.

We try to verify whether there is empirical support for the idea that adverse work incentives in social security systems might push elderly in a poverty trap. This means that people may be given an incentive not to work while at the same time they slip into poverty.

One does not deny that early retirement may be the result of unfavourable labour demand and be perceived as an involuntary forced⁵ choice by the individual, “an offer he cannot refuse”. On the contrary. But this does not impede that the individual may at the same time have an incentive not to look for a job or become demotivated and discouraged. The question to be analysed is whether duration dependence in poverty among elderly that are retiring from the labour force is true or spurious. To the extent that individual characteristics like low abilities or unfavourable attitudes persist over time, they may also be the reason that individuals persist in poverty over time. In that case, one expects that once individual characteristics are controlled for, duration dependence becomes spurious. The alternative possibility is that poverty experience has a genuine causal impact on future poverty. According to Biewen(2003) and Biewen(2004), this may be because low income may be associated with adverse work incentives which make it not worthwhile for an individual to take up a job if unemployed or even to keep a low-paid job (the poverty trap). Or this may be due to loss of motivation or depreciation of human capital which complicates the search for a new job. The reasons for dependence in poverty are obviously of interest for developing effective poverty reducing measures since true dependence would suggest to focus on stigma and adverse work incentives while

³ Fenge-Pestieau(2005), Blöndal-Scarpetta(1998), Grüber-Wise(1999), Grüber-Wise(2004).

⁴ Boldrin-Dolado-Jimeno-Peracchi(1999), Fenge-Pestieau(2005)), Jousten-Lefebvre-Perelman-Pestieau(2008).

⁵ Dorn-Souza(2005), Lindeboom(1998), Jousten (2005).

spurious dependence would suggest to change individual's characteristics and abilities. Following Meghir-Whitehouse(1997), Devicienti(2002), Biewen(2003), Fertig-Tamm(2007) and Hansen-Wahlberg(2004) the question whether duration dependence reflects true duration dependence or individual heterogeneity is analyzed through a multiple-spell model of transitions in and out of poverty, controlling for observed and unobserved individual heterogeneity and for a potential initial condition problem. For Belgium, this issue has been addressed for welfare spells in Cockx(1998) and for unemployment spells(Cockx-Dejemeppe(2005)).

This paper is structured as follows. Section 2 explains the matching between the administrative dataset and Census. Section 3 describes at the aggregate population level the observed poverty transition rates associated with the duration of a (non)poverty spell. If there are individual-specific unobserved factors that affect the hazard, the aggregate transition rates will tend to be different from those at the individual level. In order to examine this, section 4 estimates a multiple spell discrete-time hazard model through which we estimate simultaneously exit and re-entry rates while allowing for observed and unobserved heterogeneity and controlling for a potential initial condition problem. This section also contains an overview of the modelling literature on poverty dynamics. Section 5 concludes.

2. Data Construction

The results presented below are derived from a micro-dataset provided by the National Institute for Statistics that contains information on 30183 Belgian households⁶ of which at least one member is between 55 and 75 years old on 31 December 2001 and that were randomly selected out of the National Register. Of these 30183 households, one keeps

⁶ That is 60806 individuals. A household is defined as the number of individuals having the same domicile, as registered in the National Register. On the basis of the National Register of 2001, 93.66% of the 60806 individuals are head(49,6%), partner of the head(29.35%) or child(14.72%) of the head of the household. Only 6.33% has another kind of relationship to the head: 1.25% are grandchild, 0.76% parent of the head, 0.47% child-in-law, 0.48% parent in law, 0.46% brother or sister and 2.22% are no family related habitants,...

the individuals that are between 55-75 years old in 2001. This reduces the number of individuals of the final dataset from 60806 to 43726.

These 30183 households (corresponding to 43726 individuals) have in the first place been connected to the Income Tax Returns data (1991 – 2002)⁷ (=ITR) by means of the national identification number. This implies that the individuals in our dataset are between 45- 65 years old in 1992, 46-66 in 1993, 47-67 in 1994,... This administrative dataset contains the yearly information necessary to calculate the income tax of the fiscal household to which the individual belongs. The variables it includes are civil status, number and type of dependants in the fiscal household, gross capital income, the age and gross labour income, replacement incomes of the household members (old-age pension, early retirement, unemployment, illness or disability benefits, ...), housing wealth, occupational pension benefits, employee contributions in occupational pension plans, private subsidized savings. Every year, about 86% of the individuals selected out of the National Register could be matched with the ITR. This means that in a given year about 14% of the Belgian civil population are not in the ITR.⁸ However, only 4.2% of the individuals selected out of the National Register and between 55-75 years old in 2001 (that is 1844 of 43726) or 4.9% of the households (that is 1502 of 30183), do not appear in the ITR for any year 1991-2002 and will have to be deleted from the analysis altogether⁹. We will discuss this issue below. The stability over time in the percentage of individuals out of the ITR masks thus considerable income mobility at the individual level in and out of the ITR.

Interestingly the 30183 households (corresponding to 43726 individuals) selected out of the National Register could also be merged with the Census of 2001 through the use of

⁷ Note that 723 of the 30183 households group several fiscal household units. For example, two widowers/divorced/single individuals living together but filling in separate tax files are two fiscal households but will be considered as one household. However, one does not know whether before 2001 these individuals are living together what may generate measurement errors. We suppose they are living together from the first year in which they are observed both as declarant in the ITR while it may be they start living together only a few years afterwards.

⁸ This confirms the finding of Perelman-Schleiper-Stewart(1998) that 13% of the Belgian population do not declare incomes.

⁹ However, their full socio-economic profile can be established since they are in the Census of 2001.

the national identification number. This survey has a response rate at the level of the individual of 98,7%. It contains detailed information on education level, professional category (private sector employee, civil servant, self-employed,...), the sector the household member works or worked in (agriculture, banking, construction, transport, chemical industry, real estate, army, retail, ...) and also the self-reported general health status.

Table 1: Construction Dataset		
	Number of households	Number of individuals
National Register	30183	60806
National register 55-75 years old	30183	43726
National Register 55-75 years old, matched with Census of 2001	30183	43726
matched with ITR	29760 (matching: 98,6%)	43157(matching: 98,7%)
1991		
1992	25830 (matching:86,3%)	37423 (matching:85,7%)
1993	25515 (matching: 84,5%)	37023 (matching: 84,67%)
1994	25559 (matching: 84,6%)	37115 (matching: 84,88%)
1995	25326 (matching: 83,9%)	36856 (matching: 84,29%)
1996	24778 (matching: 82,0%)	36208 (matching: 82,80%)
1997	25566 (matching: 84,7%)	37213 (matching: 85,10%)
1998	25727 (matching: 85,2%)	37465 (matching: 85,68%)
1999	25850 (matching: 85,6%)	37651 (matching: 86,10%)
2000	25855 (matching: 85,6%)	37677 (matching: 86,16%)
2001	25914 (matching: 85,8%)	37764 (matching: 86,36%)
2002	25982 (matching: 86,0%)	37819 (matching: 86,49%)
	26269 (matching: 87,0%)	38213 (matching: 87,39%)
Number of individuals (households) that never appear in ITR between 1991-2002: 1844 (1502 ¹⁰)		
Number of individuals(households) that appear at least one year in ITR between 1991-2002: 41882 (28681 ¹¹)		

As can be seen from table 1, the number of households is not for every year of the ITR the same: households may temporally or permanently drop out of the ITR. According to article 178 of the Royal Decree corresponding to the Belgian income tax code of 1992,

¹⁰ 1635 households with a member that does not declare - 133 households where at least one member declares income.

¹¹ 28548 households where every member declares+133 households where at least one member declares income.

are not obliged to declare incomes: 1° households without professional activity with an income below the minimum taxable income (except singles/widow(er)s with dependent children) and 2° households of which the income only consists of old-age pensions and housing wealth. It is unfortunately¹² impossible to know whether a dropout is due to 1° or 2°¹³. Individuals that drop out because of reason 1° may from the moment their income exceeds the minimum taxable income reappear in the ITR. If that happens we qualify missing periods before the reappearance as periods in poverty. Eligibility rules for old-age pension benefits can also be used to qualify missing observations as periods in poverty. Table 2 shows how this correction reduces the number of unbalanced households. Individuals that drop out because of reason 2° or because of reason 1° that do not reappear, will from that moment on never appear again in the ITR. We treat these drop-outs as right-censored observations¹⁴.

Table 2: Unbalanced data				
Number of years observed	Number households after correction	% after correction	Number households before correction	% before correction
1	279	0.97	671	2.34
2	199	0.69	469	1.64
3	217	0.76	499	1.74
4	229	0.80	476	1.66
5	277	0.97	567	1.98
6	266	0.93	548	1.91
7	359	1.25	651	2.27
8	380	1.32	689	2.40
9	481	1.68	825	2.88
10	453	1.58	1169	4.08
11	845	2.95	2743	9.56
12	24696	86.11	19375	67.55
<i>TOTAL</i>	<i>28681</i>	<i>100.00</i>	<i>28681</i>	<i>100.00</i>

¹² A technical possibility was to merge the data with an existing dataset of means-tested beneficiaries for the years 1991-2001 but access to this dataset was not allowed, although this would enable us to identify (part of) 1°.

¹³ For sure, a dropout cannot be due to death since the sampling scheme also implies that individuals are sampled conditional on being alive in 2001. If poor singles would be more likely to die this could induce an endogenous selection problem. Up to this stage, this has been neglected.

¹⁴ As Devicienti(2002), Stevens(1999), Devicienti-Gualtieri(2007), Fertig-Tamm(2007). However to the extent that the dropouts because of reason 1° cannot be considered as random censoring, the sample selection problem they might induce should be modeled explicitly. This is an issue for future research. Capellari-Jenkins(2004) models attrition simultaneously with poverty transitions for UK and Lillard-Panis(1998) models household composition and attrition simultaneously for US but both find that attrition induces a negligible bias in the estimation results.

The correction reduces the number of missing observations of individuals that appear at least once in the ITR from 11.9% to 5.4% of total observations¹⁵. The effect of this correction is to increase poverty persistence, increase the number of households that is once confronted with poverty and to change the coefficients of duration dependence in the regressions. However it doesn't change any of the general conclusions on the issues under study.

Finally we calculate for every household¹⁶ its net income, convert it in real terms¹⁷ and inflate net real income by the OECD equivalence scale that attributes 1 to the head of the household, 0.5 per additional adult and 0.3 per child.

In the remainder, the focus of attention will be the head of the household¹⁸ but the unit for calculating income is the household. This means the head of the household is qualified poor if the income of the household to which he pertains is below the poverty line. The latter is defined as 50% of median net equivalised income¹⁹ of the whole economy.

3. Descriptive analysis of the dynamics of poverty

The longitudinal dataset presented in the previous section is well suited to describe the dynamics of poverty. We will start with a description of flows into and out of poverty and the distribution of periods spent in poverty.

¹⁵ The number of missing observations (41190 out of a total of 344172 (=12 years times 28681 households that appear at least once in the ITR observations) could be reduced to 18888. We also qualified the missing observations (1792 = 4.3% of 41190) of households with 0 years in poverty, at least 10 years out of poverty and mean income pooled over all available years of more than 140% of the poverty line as non-poor. These correspond to measurement errors of old-age pensioners or temporary emigration but not to a poverty experience.

¹⁶ For two widowers or divorced living together, net incomes are calculated separately for each fiscal household. Then the net income of all fiscal households is summed and in a last step the equivalence scale is applied on this sum.

¹⁷ With year 2002 as reference year.

¹⁸ The head of the household is the individual that declares income. For married individuals, the fiscal legislation says it is the man. In 2838 households the spouse of the declarant becomes declarant herself due to death of her husband or divorce. Similarly, if the head of the fiscal household is not between 55-75 years old in 2001, but the partner of the head is, we take the partner as head of the household (1983 households). In cases with two fiscal households living together (723 households), the head is the member that is most years in the ITR.

¹⁹ That were kindly provided by the National Institute for Statistics.

3.1. To start, table 3 shows the distribution of the total number of years spent in poverty.

Table 3: distribution of total number of years spent in poverty²⁰			
Number of years in poverty	Number of households	% of those who are poor at least once	% of population
0	17997	/	0.63
1	2435	0.22	0.085
2	1212	0.11	0.042
3	944	0.086	0.032
4	721	0.066	0.025
5	781	0.071	0.027
6	624	0.057	0.021
7	586	0.053	0.020
8	513	0.047	0.017
9	548	0.050	0.019
10	577	0.053	0.020
11	644	0.059	0.022
12 or more	1099	0.101	0.038
Number of households that are at least 1 year in poverty: 10684			
Number of households: 28681			

The fact that 63% of the elderly households is never poor implies that 37% of them is once confronted with poverty. This is much higher than the “static” poverty rates of around 12% that are found on the basis of the PSBH in Cantillon(1999) and also on the basis of our own dataset. Deleeck-Cantillon(1992) find similarly on the basis of two waves of the SEP that of the whole population 10.8% is poor²¹ in 1988 and 1985, 73% is not poor in 1988 and neither in 1985 while 16.2% is once poor during that period.

For our data, the transitory poor, that are poor for exactly one year, account for 8.5% of all households. Those that are poor for at least 3 years make up 66% of those who ever have been poor and 24% of all households. In general most people that slip into poverty are quite successful in getting out. But precisely because this is true the transitory poor

²⁰ Including left-censored spells

²¹ Poverty line is 50% of average income.

are a small fraction of the poor at any point in time and those with longer poverty spells account for the bulk of all poverty.

3.2. The degree of persistence of poverty and the recurrent poor cannot be read from table 19. It is not because households are in total 3 years poor that these are 3 consecutive years in poverty. The persistently poor are poor for at least 3 consecutive years. The recurrent poor are poor for at least one year but never longer than 2 consecutive years. It is thus possible that a recurrent poor is for example poor for 5 years in total but is not persistently poor. Table 4 shows that more than 60% of the elderly who once have been poor are persistently poor.

Table 4: Persistent and recurrent poverty among elderly²²			
	Number of households	% of population	% of those who are poor at least once
Persistent	6499	0.22	0.61
Recurrent poor	4185	0.14	0.39
Number of households that are at least one year in poverty: 10684			
Number of households: 28681			

3.3. The number of consecutive years one is in poverty defines a poverty spell. When studying poverty spells arises the issue of censored spells. Suppose a household is counted poor for exactly one year. We would qualify it as transitory poor. However if that year corresponds to the first/last year of observation the duration of the poverty spell is underestimated if the household was poor before the sampling began/after the sampling stopped. The following table displays the considerable percentage of censored poverty spells. It will be discussed below how account will be taken of censored spells.

Table 5: Censored poverty spells			
	Number of households	% of population	% of those who are poor at least once
Left-censored	5082	0.17	0.47
Right-censored	5200	0.18	0.48
Left and right censored	2758	0.09	0.25
Number of households that are poor at least once: 10684			
Number of households: 28681			

²² Including left-censored spells.

3.4. From the moment that the data are arranged such that each household is associated with the duration of one (or more) spells, one can calculate exit and (re)entry rates. The exit rate associated with a given duration of a spell is the number of households that exit at that length of the spell divided by the population at risk of exiting. The survivor function associated with a certain spell duration specifies the probability that an individual will survive in that spell beyond that duration. Spells that are right-censored are included in all but the censored year. Figures 1 and 2 plot the poverty exit rates and survivor function with and without left-censored spells but as can be seen they do not differ a lot from each other²³.

Figure 1: Survivor function with and without left-censored spells

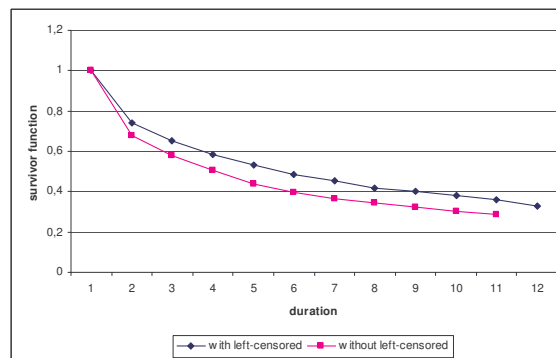
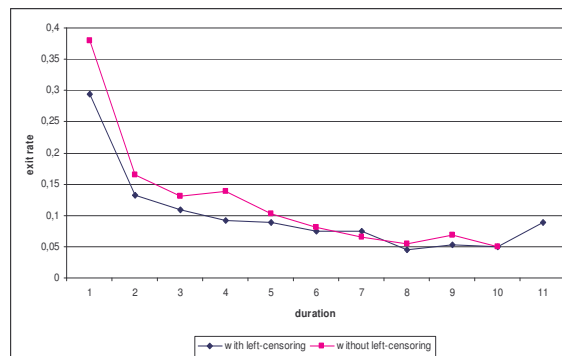


Figure 2: Exit probability with and without left-censored spells



The exit rate is high for the transitory poor and low for those that are long-term poor. The survivor function decreases sharply in the first years but after some years it seems to

²³ The exclusion of left-censored spells implies that exit rates can only be calculated for a duration of maximum 10 years.

stabilize. This means that the longer somebody is poor, the more difficult it becomes to leave poverty.

To find the (re)entry probabilities, one calculates for each of the households at risk of (re)entering poverty the length of the spell that they are out of poverty. Then for each possible length of the spell the number of individuals that enter poverty is divided by the population at risk of (re)entering. The results, whether left-censored spells are included or not, are in figures 3 and 4 and differ quite a lot²⁴. The reentry rates with left-censored spells are commonly called entry rates. Probabilities of entering poverty are very low and around 1-2%: they are based on households that may or may not have been poor once. Re-entry rates are up to 4-6 times higher than the entry rates indicating that the probability of becoming poor is much higher for households who have been poor than for those who have not.

Figure 3: (Re)entry probability with and without left-censored spells

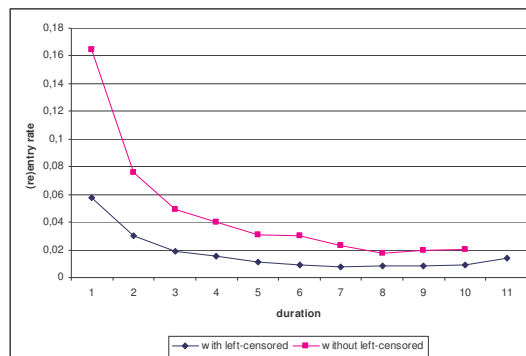
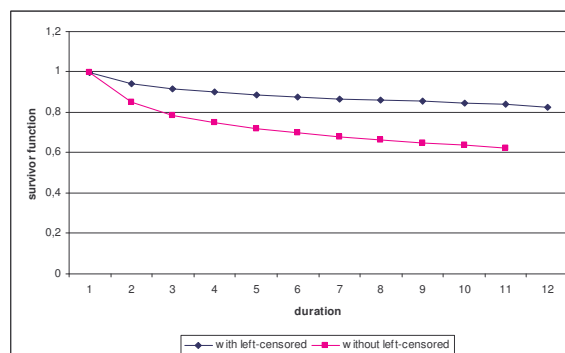


Figure 4: Survivor function with and without left-censored spells



²⁴ The exclusion of left-censored spells implies that the entry rates can only be calculated for a duration of maximum 10 years.

Survivor functions represent the cumulative risk of slipping back into poverty after a previous exit. Of those who quit poverty, about 60% are still out of poverty after 12 years. Thus 40% of those who quit poverty will fall back into poverty.

The question that interests us is whether the exit rates out of poverty and reentry rates into poverty, that are downward sloping, remain so, after one has controlled for all kind of individual effects. To the extent that individual characteristics like the lack of abilities persist over time, they may be the reason that Belgian elderly persist in poverty over time. If this would be true, one expects that after controlling for individual-specific effects, duration dependence will no more be significant. Alternatively, if current poverty experience has a causal impact on future poverty experience, one expects that negative duration dependence remains significant. To check this, a multivariate analysis is necessary, to which we turn now.

4. A multiple-spell discrete-time hazard model of poverty dynamics

The purpose of this section is to verify whether duration dependence in the exit and re-entry rates is due to individual heterogeneity or true duration dependence through the estimation of a multiple-spell discrete-time hazard model while controlling for unobserved effects and a potential initial condition problem (4.3.). We start this section with an overview of the existing empirical models of poverty dynamics (4.1.), followed by a presentation of the model that will be used for our purposes (4.2.).

4.1. Previous modelling research on poverty dynamics

We briefly overview the empirical literature because we want to point out why we do not use some type of models and why only one type of model is appropriate for our purpose. In the end, we discuss the few models that have been estimated on Belgian data.

4.1.1. Component of variance models

One of the first to study poverty dynamics are Lillard-Willis(1978) who estimate an earnings model with log of earnings y_{it} as dependent variable, individual $i = 1, \dots, N$ at time $t = 1, \dots, T$ and x_{it} a vector of observed explanatory variables:

$$y_{it} = x_{it}\beta + v_{it}$$

The error structure has the form: $v_{it} = c_i + \gamma y_{it-1} + u_{it}$ with c_i an unobserved effect, u_{it} a random error term and γ a serial correlation coefficient common to all individuals. However, if the serial correlation in the error structure results from misspecification of the population model: $y_{it} = x_{it}\beta + \rho y_{it-1} + c_i + u_{it}$, c_i is correlated with y_{it-1} and the use of instrumental variables may be necessary. Up to now, these models did not address either the fact that explanatory variables such as household composition or labour market status that are often included might be endogenous to the dynamics of income. In addition, a common variance structure is assumed for the entire population while the dynamics of individuals in different parts of the income distribution might be different. According to Stevens(1999) and Devicienti(2002), they perform less than duration models in predicting poverty. A recent application is Fouarge-Muffels(2003).

4.1.2. Duration models and transition probability models

Bane-Ellwood(1986)²⁵ calculate duration dependent exit probabilities and the distribution of entering poverty and of exiting poverty as a function of events (change in household composition or household income). They also calculate the expected duration of poverty as a function of events associated with the beginning of a poverty spell. Although this is not a multivariate analysis, they initiate a new strand of a literature that analyzes the determinants of flows into and out of poverty.

²⁵ For the US. For the UK, see Jenkins(1998), Jenkins(2000), Jenkins-Rigg(2000). For Germany, see Otto-Siedler(2003). For Germany, US, UK, Canada, see Antolin-Dang-Oxley(1999).

Transition probability models are developed with probability of (re)entry in and exit out of poverty as dependent variable and as independent variables change in employment status and/or household composition, components of household income, individual and household characteristics²⁶. At the same time models arise that estimate the probability of (re)entry in and exit out of poverty as dependent variable with dummies for duration, individual and household characteristics as independent variables. The latter are in fact equivalent to discrete-time hazard models, as shown by Allison(1984) and Jenkins(1995).

More complicated discrete-time hazard models extend the analysis from single to multiple spells. Most of these estimate poverty transition equations separately²⁷ under the hypothesis that, for a given individual, entry rates and exit rates can be treated as conditionally independent and that multiple spells of the same event type are conditionally independent. A first extension to this, as Callens-Croux-Avramov(2003) and Arranz-Cantto(2006), is to allow for different baseline hazards in case of multiple spells of the same event type for a given individual. An additional extension captures not only correlation between spells of the same event type but in addition that individuals with high(low) exit rates have lower(higher) reentry rates. Stevens(1999)²⁸ is the first to estimate exit and reentry equations of poverty simultaneously while allowing the unobserved effect of the two transition equations to be correlated. This leads to more accurate estimates of total time spent in poverty. The same method of joint estimation of entry and exit equations is applied by Devicienti(2002), Jenkins-Rigg (2001), Biewen(2003), Fertig-Tamm(2007), Devicienti-Gualtieri(2007), Biewen(2003), Wahlberg-Hansen(2004). For these joint estimations, the most frequent distributions for

²⁶ Bourreau-Dubois-Jeandidier-Berger(2003), Bardasi-Jenkins-Rigg(2000), Valletta(2005), Dewilde(2004), McKernan-Ratcliffe(2002), Antolin-Dang-Oxley(1999), Fouarge-Layte(2003).

²⁷ McKernan-Ratcliffe(2002), Finnie-Sweetman(2002), Devicienti(2001a), Antolin-Oxley-Dang(1999), Zoyem(2002), Fouarge-Layte(2003), Valletta(2005), Dewilde(2004), Makovec(2005), Bourreau-Dubois-Jeandidier-Berger(2003), Devicienti(2007), Capellari(2007), Jenkins-Rigg(2001), Arranz-Cantto(2006). Arranz-Cantto(2006), McKernan-Ratcliffe(2002), Finnie-Sweetman(2002), Antolin-Dang-Oxley(1999), Fouarge-Layte(2003) and Devicienti(2001a) include however the number or length of previous spells, although this is an endogenous variable. Note also that even if the exit and entry processes would be independent, estimating them jointly instead of separately would be more efficient.

²⁸ Following Meghir-Whitehouse(1997) who estimate jointly unemployment and employment spells, while accounting for unobserved effects and an initial conditions problem.

the unobserved effect are the discrete distribution with a finite number of support points²⁹ and the multivariate normal distribution³⁰.

There exist also discrete-time duration models that estimate exit and reentry rates separately while accounting for unobserved heterogeneity: Capellari(2007), Finnie-Ross(2002), Fouarge-Layte(2003), Makovec(2005). However, it appears³¹ that for single spell data, estimation results are sensitive to misspecification of the distribution of unobserved effects while for multiple-spell models, it is much more easier to estimate parameters that are robust to the functional form of the unobserved effect³².

The discrete-time hazard model takes account of right-censored spells³³ under the assumption they are randomly censored but left-censored spells are more problematic. Excluding them, as a lot of models do³⁴, could result in a sample selection bias in the presence of unobserved heterogeneity. Lancaster(1990) notes: “The common treatment of stock sampled data with future spells observed is to ignore the elapsed duration data and to base inferences solely on those spells that begin after the sampling data. This is a sensible and correct way to proceed in models that do not involve unmeasured person-specific heterogeneity. Unfortunately, in models that do involve such heterogeneity there is a further complication to consider due to the fact that the distribution of unobservable quantities also depends of the sampling scheme”³⁵. If the probability that the first spell will be poverty or non-poverty depends on individual characteristics including any unobserved heterogeneity, excluding left-censored spells creates an initial conditions

²⁹ Meghir-Whitehouse(1997), Stevens(1999),Devicienti(2002), Hansen-Wahlberg(2004), Biewen(2003), Fertig-Tamm(2007) use the Heckman-Singer(1984) estimator.

³⁰ Lillard-Panis(1998).

³¹ Heckman-Singer(1984). Meyer(1990, p.771) notes that “it is plausible that much of the parameter instability found by Heckman-Singer(1984) is due to their assumption of a Weibull baseline hazard. When the baseline hazard is nonparametrically estimated, the choice of heterogeneity distribution may be unimportant”. Nicoletti(2006,p.19) finds in the same line for discrete-time hazard models that “misspecifying the random effects distribution biases neither the duration dependence nor the covariate coefficients estimation”.

³² van den Berg(2001).

³³ Each individual contributes to the likelihood function for every year he is in the dataset but depending on whether the last interval is censored or not the dependent variable of the last contribution equals 0 or 1.

³⁴ Bane-Elwood(1986), Antolin-Dang-Oxley(1999), Finnie-Sweetman(2002), Devicienti(2001a - 2002), Fouarge-Layte(2003), Makovec(2005), Devicienti(2007), Arranz-Canto(2006). The problem of left-censoring is absent in the case of flow sampling (Cockx(1998),Dejemeppe-Cockx(2005)).

³⁵ P.189.

problem whereby the identity of the first spell is endogenous. To control for this selection bias, Devicienti(2002), Biewen(2003), Fertig-Tamm(2007), Wahlberg-Hansen(2004) extend the analysis of Stevens(1999) and estimate the entry and exit equations jointly with an initial condition equation, as Meghir-Whitehouse(1997), while allowing the unobserved effects in the transition equations to be correlated with the unobserved effect in the initial condition equation. In order to identify the model, the initial condition equation contains exclusion restrictions. Whitehouse-Meghir(1997) use unemployment rates at the first time the spell is observed, Fertig-Tamm(2007) dummies for the year of the first observation and education level of the parents of the household head, Biewen(2003) the education level of the parents and city where the individual grew up, Devicienti(2002) the education level of the parents of the household and Hansen-Wahlberg(2004) uses no exclusion restrictions at all what implies that the model is supposed to be identified by the functional form of the unobserved effects. Another possibility to treat the sample selection bias induced by stock-sampling is to include left-censored spells but to correct the likelihood function for the fact that the length of the first spells is underestimated. This requires the assumption of a constant entry rate and a constant survivor function or additional data such as the aggregate number of entering poverty at each calendar date in the past as is empirically relevant such that one can construct, as Nickell(1979), a model for the entry rate and survival function. There are no such data for Belgium and thus one cannot apply this method.

4.1.3. Dynamic unobserved effect models

Dynamic unobserved effect models are developed in unemployment dynamics literature by Heckman(1981), Arulampalam-Booth-Taylor(2000), Stewart-Sheffield(1999), Stewart(2007) and also have found poverty applications. Specify the model for individual $i=1,\dots,N$ at time $t=2,\dots,T$ as

$$y_{it}^* = x_{it}'\beta + \gamma_{it-1} + c_i + u_{it}$$

Where y_{it}^* denotes the unobservable propensity to be poor, y_{it-1} the observed poverty status in t-1, x_{it} a vector of observable characteristics, c_i an unobserved effect and u_{it} a random error term. y_{it} is the dependent variable

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{else} \end{cases} . \text{ The inclusion of the lagged dependent variable allows to test for}$$

state dependence. In contrast to discrete-time hazard and components-of-variance models that take into account a history of lags, this model assumes thus that one lag of poverty status is sufficient to capture its full dynamics. According to Devicienti-Gualtieri(2007) the significant duration dependence in observed transition rates casts doubts on the first-order lag often assumed in empirical work. To estimate this model, one first integrates out the unobserved effect by assuming a distribution for c_i that is independent from x_{it} , usually a normal or Gamma. Secondly if the initial observation of y_{it} is correlated with c_i , this raises the initial conditions problem. According to Heckman(1981), Cappellari-Dorsett-Haile(2007) estimate jointly with the poverty transition equation an initial condition equation while allowing for correlation between the unobserved effects affecting the poverty transition and initial condition equations. With the idea that initial labour market states, but not transitions, depend upon the macroeconomic conditions prevailing at the time, they use GDP growth rate measured in the year of the first observation as exclusion restriction. Another approach has been suggested by Wooldridge(2002) that consists of modelling the distribution of the unobserved effect conditional on the initial poverty status. It is applied by Poggi(2007) and Biewen(2004)³⁶. Biewen(2004) challenges in addition the assumption of strict exogeneity of the explanatory variables, that is that there must not be any feedback from current poverty to future values of the explanatory variables employment status and household composition. To this end, he estimates a joint dynamic random effects model of poverty status, employment status and household composition status. Since the

³⁶ McKernan-Ratcliffe(2002) condition the unobserved effect on the initial poverty status in a discrete-time hazard model. They find that individuals with left-censored (non)poverty spells are significantly less likely to (enter)exit poverty.

assumption of one lag of poverty status is unreasonable in our case, we did not estimate this model.

4.1.4. Markovian transition models

Capellari-Jenkins(2002), Capellari-Jenkins(2004), Cappellari(1999), VanKerm(2004) model entry and exit probabilities simultaneously using an endogenous switching regression model with a binary dependent variable³⁷. Specify the poverty transition equation for individual $i=1, \dots, N$ at time $t=2, \dots, T$ as

$$y_{it}^* = [(y_{it-1})\gamma_1' + (1 - y_{it-1})\gamma_2']x_{it-1} + c_i + u_{it}$$

With y_{it}^* the unobserved propensity of being poor in t , y_{it-1} poverty status in $t-1$, γ_1' and γ_2' the coefficient estimates conditioning on being poor respectively non-poor in $t-1$, x_{it-1} a vector of observable characteristics, c_i an unobserved effect and u_{it} a random error term. y_{it} is the dependent variable

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{else} \end{cases} . \text{ These models typically account for different sources of non-random}$$

selection such as attrition and initial conditions by estimating jointly an initial condition equation and attrition equation with the poverty transition equation while allowing the unobserved effects of these equations to be correlated with a multivariate normal distribution. Heckman(1981) suggested the use of pre-sample information such as the education level, occupation, labour market status of the parents of the household head as exclusion restriction. Model estimates can be used to derive predictions of the poverty persistence rate and entry rate. In contrast to the dynamic unobserved effect models, the lag structure in x_{it-1} rules out the possibility of instantaneous effects of changes in characteristics for poverty status. For example changes in employment status are not allowed to affect poverty until the next period. In addition, one may have doubts on the

³⁷ Following Stewart-Swaffield(1999) who used this model to estimate low pay dynamics.

appropriateness of the first-order dynamics assumption³⁸. That is why we did not use this model.

4.1.5. Structural models

Aasve-Burgess-Dickson-Propper(2006) argue that poverty is not a decision variable but rather the outcome of underlying behavioural decisions such as whether to work, to have children, to marry and divorce. They estimate 5 simultaneous hazards (childbearing, marriage, divorce, employment and non-employment) while allowing the unobserved effects of these 5 equations to be correlated according to a multivariate normal distribution. From these results, one can derive a model for income dynamics and poverty. In their model all persistence within poverty is attributed to persistence within demographic and labour market states. Although already very time-consuming and complex, they only account for two labour market statuses such as employed versus non-employed while we are interested in the transitions between disability, unemployed, self-employed or employed, old-age pension and early retired. In addition there is no consensus on whether these labour market transitions are voluntary or involuntary: early retirement may result from an unfavourable labour demand. Finally, changes in household composition, like becoming widow, can difficultly be treated as a behavioural decision. For these reasons, we did not estimate this model.

After this overview of the existing research on modelling poverty transitions, let us look what models have been applied to Belgian data. In Belgium, models on the duration of unemployment spells (Dejemeppe-Cockx(2005)) and welfare spells(Cockx(1998)) are available, as well as models on poverty persistence but not for elderly. Van Kerm(2004) finds, through the estimation of a Markov transition model, on the PSBH on a population between 25 and 55 years old that poverty entry depends on household and employment status. In particular being unemployed, self-employed or single increases the risk of poverty entry. He controls for the endogeneity of the initial poverty status by estimating jointly with the poverty transition equation an initial condition equation while allowing

³⁸ Devicienti-Gualtieri(2007).

the unobserved effects of these equations to be correlated with a trivariate normal distribution. He does not consider the possibility that household and employment status that are included as explanatory variables might be endogenous. Indeed, current poverty status may affect future employment and household composition. Similarly, Nicaise-Deblander(2005) estimated a Markovian switching model on the PSBH for the years 1993-1997 for the working age population, while controlling for initial condition equations. As an extension to VanKerm(2004), Nicaise-Deblander(2005) controlled in addition for possible endogeneity of employment status (but not for household status). They find that initial employment status is insignificant to explain initial poverty status and transitions into poverty but is significant in explaining poverty exit³⁹. Household composition has a significant effect on poverty transitions. Dewilde (2004) finds through the separate estimation of a transition probability model on the PSBH for the whole population that poverty entry and exit depend on household and employment related events. In particular entry into unemployment, disability, (early)retirement is associated with entry into poverty. Makovec(2005) uses a discrete-time hazard model to estimate separately poverty entry and exit equations for those above 55 years old in the ECHP, taking into account unobserved heterogeneity. For the entry model, he finds that accounting for unobserved effects leads to spurious dependence. Like Vankerm(2004), Makovec(2005) includes household and employment status as explanatory variables, ignoring possible endogeneity. He includes dummies for employed versus non-employed, for aged above or below 65, for receipt of disability and old age benefits. None of the above dummies is significant except the receipt of disability benefits that increase poverty exit. That is the opposite result of Dewilde(2004) although the dataset in the ECHP for Belgium is also the PSBH. Dewilde(2004) though does not control for duration effects, nor for unobserved effects that are highly significant in Makovec(2005). On the other hand, Makovec(2005) used a very small sample limited to those above 55 years old while also Dewilde(2004) reports problems with limited sample size. Although he conditions on unobserved effects, Makovec(2005) does not adress the issue of left-censored spells.

³⁹ Instead of a simultaneous estimation of poverty and employment equation, they followed a less efficient two-step procedure.

4.2. The model

Following Meghir-Whitehouse(1997), Stevens(1999), Devicienti(2002) and Biewen(2003), we estimate the discrete-time hazard model, given our interest in duration dependence. Although in the real world poverty transitions can occur at any time, the model is in discrete time since the data are grouped into intervals of one calendar year. There are two types of spells: poverty spells ($k=p$) and non-poverty spells ($k=np$). We assume that the probability that an individual $i=1, \dots, N$ leaves the spell of type k in the calendar year $t=1, \dots, T$ at a duration d is defined as $P_{it}^k = \text{Prob}(y_{it}^{*k} > 0)$ that results from the latent model $y_{it}^{*k} = c_i^k + f^k(d) + \beta^k x_{it} + u_{it}^k$ where y_{it}^{*k} denotes the unobservable propensity to be in a spell of type k , x_{it} is a vector of observable characteristics, β^k the vector of coefficients associated with x_{it} , $f^k(d)$ is a function of duration dependence that represents the baseline hazard and where $d=1, \dots, D$ denotes the duration of the current spell and D is the maximum duration of a spell. We will adopt a flexible specification for the baseline hazard⁴⁰ where $f^k(d) = \alpha_1^k DU_1 + \alpha_2^k DU_2 + \dots + \alpha_D^k DU_D$ and DU_d are dummies corresponding to a duration d . We assume that individuals enter a spell at $d=0$ and are at risk of leaving the spell at $d=1, \dots, D$ ⁴¹. Unobserved heterogeneity enters the specification of the hazard rate as an individual-specific additive error term c_i^k constant over time which is allowed to be correlated across different types of spells. We are thus estimating a binary response model where the dependent variable

⁴⁰ Lancaster notes “there is nothing to be gained by adopting the more elaborate proportional hazard model over the piecewise-constant one. Essentially, since with grouped data we can know nothing about the way the hazard varies within the interval, the best we can do is to estimate its average level and we might as well work with the simplest model, in which that level is constant”, p.181.

⁴¹ Wooldridge(2002) notes: “usually the duration dummies are unrestricted, in which case x_{it} does not contain an intercept” (p.709). Alternatively, as explained by Jenkins(2001), one can drop one duration dummy to use it as a reference and fit an overall intercept term of the model. In our 12 year data set, the exclusion of left-censored spells implies that exit rates can only be calculated up to a maximum duration of 10 years (see also figures 2 and 3). So $D=10$ if we do not use an intercept and $D=9$ if we do use an intercept term.

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{else} \end{cases} \quad \text{and } u_{it}^k \text{ is a random error term. If } u_{it}^k \text{ is logistically}^{42} \text{ distributed,}$$

$$P_{it}^k = \frac{\exp(c_i^k + f^k(d) + \beta^k x_{it})}{1 + \exp(c_i^k + f^k(d) + \beta^k x_{it})}. \quad \text{In the absence of unobserved heterogeneity } (c_i^k = 0)$$

one assumes that the exit and re-entry equations represent conditionally independent processes and therefore the log-likelihood function can be maximised separately for all

$$\text{spells of a given type } k \text{ as: } \log L^k = \sum_{i=1}^N \sum_{t=1}^T m_{it} \{ (1 - l_{it}) \log(1 - P_{it}^k) + l_{it} \log(P_{it}^k) \} \text{ where } l_{it}$$

indicates whether an exit from the spell is observed for individual i in t ($l_{it} = 1$) or not ($l_{it} = 0$) and $m_{it} = 1$ if for individual i in t a spell of type k is being observed and $m_{it} = 0$ otherwise.

If the transition equations depend on a random effect that is allowed to be correlated across spells of different types ($c_i^k \neq 0$), poverty and non-poverty spells cannot be treated separately and a simultaneous estimation is necessary. One might expect indeed that there are individual-specific unobserved effects like ability, motivation or general attitudes that affect each type of transitions. If individuals have a high propensity to leave poverty, one may expect they have also have a low propensity to reenter poverty. If that would be so, there would be negative correlation among the unobserved effects of

⁴² Studies on poverty dynamics use a logistic, normal or extreme value distribution for the random error term. Stevens(1999) and Devicienti(2002) use a logistic distribution for the random error term, Meghir-Whitehouse(1997), Fertig-Tamm(2007), Biewen(2003) and Hansen-Wahlberg(2004) a normal distribution, Bardasi-Jenkins-Rigg(2001) an extreme value distribution. Sueyoshi(1995) explores the implications of these specifications for hazard behaviour and notes that “practical experience with discrete-choice models suggests that the predicted probabilities and hence the goodness-of-fit tests for the models will generally be quite similar.” Apart from goodness-of-fit tests, “results from the logit and proportional hazard specifications will be quite similar. In contrast, estimates from a probit-type group duration model should depart significantly from both of these specifications, exhibiting covariate effects that are decidedly non-proportional” while “logistic models are only slightly less proportional than the extreme value specification”. The fact that extreme value and logistic estimation give very similar results is because, if h denotes the hazard rate, the odds ratios (the exponentiated coefficients of the logistic model) will tend, if the hazard is sufficiently small, to the hazard ratio where the latter corresponds to the exponentiated coefficients of the extreme value model. To check whether the coefficients of the extreme value and logistic model differ a lot, we estimated the model again with an extreme value distribution of the error terms. However, coefficients and standard errors are very similar. The predicted hazard rates are barely distinguishable.

the transition equations. In a model that allows for unobserved heterogeneity, an additional problem arises. The probability that the first non-left censored spell is a poverty spell will depend on individual characteristics including unobserved effects, creating an initial conditions problem whereby the identity of the first complete spell we observe is endogenous. To control for the selection bias that may arise we follow Heckmans(1981) approximation method and define a probability of being in a spell of type k at the initial year of observation as a function of individual characteristics and unobserved effects and we estimate this probability together with the transition equations while allowing the unobserved effect of the initial condition equation to be correlated with the unobserved effects of the transition equations. To identify the model we use explanatory variables in the initial condition equation that are excluded from the transition equations:

$$P_i^0 = \frac{\exp(q + \mathcal{W}_{i0})}{1 + \exp(q + \mathcal{W}_{i0})} \text{ where } t=0 \text{ is used to denote the calendar year in which the first}$$

non-left censored poverty spell started, $W_{i0} (\neq x_{it})$ is a vector of observable characteristics and q corresponds to the unobserved effect. If we denote the joint trivariate distribution of the random unobserved effects by $F(c^p, c^{np}, q)$, the log-likelihood function for the whole sample becomes:

$$\log L = \sum_{i=1}^N \log \left\{ \int_{R(c^{np})} \int_{R(c^p)} \int_{R(q)} L_i(c^p, c^{np}, q) dF(c^p, c^{np}, q) \right\} \text{ where } \text{individuals } i$$

contribution

equals

$$L_i(c^p, c^{np}, q) = P_{i0}(q)^{np_{i0}} (1 - P_{i0}(q))^{p_{i0}} * \prod_{t=1}^T \left\{ (1 - P_{it}^{np}(c^{np}))^{1-l_{it}} (P_{it}^{np}(c^{np}))^{l_{it}} \right\}^{np_{it}} \left\{ (1 - P_{it}^p(c^p))^{1-l_{it}} (P_{it}^p(c^p))^{l_{it}} \right\}^{p_{it}}$$

where $p_{it} = 1$ when the spell in t is a poverty spell (and $p_{it} = 0$ otherwise) and if it is a non-poverty spell, $np_{it} = 1$ (and $np_{it} = 0$ otherwise). P_{0i}^p denotes the probability of observing a poverty spell in the first year of the observation. In order that the sample likelihood can be estimated⁴³ a trivariate normal distribution with mean 0 is assumed for the unobserved effects.

⁴³ The estimation requires nonlinear optimization methods. The Newton-Raphson technique we used (within the PROC GLIMMIX procedure of SAS 9.2.) is a numerical algorithm to find the first-order and second order derivatives of a log-likelihood function.

One can test the ignorability of the initial condition and of individual-specific time-constant effects on the basis of the correlations of the cross-equations error terms. We write the covariances between the unobserved effects as:

$$\rho_1 = \text{cov}(c_i^p, c_i^{np})$$

$$\rho_2 = \text{cov}(c_i^{np}, q)$$

$$\rho_3 = \text{cov}(c_i^p, q)$$

Where ρ_1 summarizes the association between unobserved effects determining poverty exit and poverty re-entry, ρ_2 the association between unobserved effects determining poverty re-entry and initial poverty status and ρ_3 the association between unobserved effects determining poverty exit and initial poverty status. If these associations are significant, there is evidence of unobserved heterogeneity and of an initial condition problem.

4.3. Results

The first column of table 6 reports the estimates for separately estimated poverty transitions that do not control for unobserved heterogeneity nor for an initial condition problem. The second column resumes the estimates from the joint estimation of exit and re-entry rates allowing for correlated unobserved heterogeneity but not for the initial condition. Including the latter as well leads to the estimates in the third column.

When one wants to control for unobserved heterogeneity and explanatory variables may be time-varying, it is difficult to relax the requirement they should be strictly exogenous. Since we only are only interested in the coefficients of duration dependence, while controlling for all kind of observed and unobserved effects, we could not include all variables at our disposal and only included explanatory variables that can be justified as strictly exogenous such as age, gender, year⁴⁴. We also consider education level

⁴⁴AGE and YEAR are time-varying but can be treated essentially in the same way as time-invariant ones, as explained by Lancaster(1990, p.21).

(measured as primary, secondary or high school education) as exogenous for individuals that at the moment they are sampled approach retirement. Being member of the second pillar is considered exogenous since in Belgium this decision is taken at latest at the age of 25 or at the start of the employment relationship and it is taken by the employer and external to the employed worker⁴⁵. Finally, there are time-varying variables that may be exogenous or endogenous like employment status or household composition. We tested for their exogeneity by regressing employment status and household composition on lagged poverty status. Since this was significant, we conclude these are endogenous and excluded them⁴⁶. The right approach would be to estimate the two poverty transitions equations (and the initial condition equation) together with an equation for employment status and marital status⁴⁷ but this is an issue for future research.

A practical difficulty concerns the choice of exclusion restrictions in the third column. The theoretical idea is clear: good exclusion restrictions should only affect the probability that the first spell of an individual that is selected is a poverty spell while it may have no effect on poverty transitions. The empirical literature does however not provide a lot of guidance on this matter. If a variable is used in the initial condition equation while it is excluded by the researcher from the transition equation although it would be significant in the transition equation, it would lead to measurement error. Therefore, we tested explicitly whether the variables that were introduced in the initial condition equation were insignificant in the transition equation. We dispose of the variable subjective reported health status that, if it could argued to be time-invariant, could be included as strictly exogenous instrument since it appeared to be strongly significant in the initial condition equation but not in the transition equations. The common practice of all models on poverty dynamics that account for unobserved effects and dispose of a health

⁴⁵ Since the law of 6th april 1995, the decision to introduce an occupational pension scheme is the exclusive authority of the employer. In addition, the age at which the employee becomes member of a scheme is at maximum 25 years old or at the start of the employment relationship.

⁴⁶ In contrast to VanKerm(2004), Makovec(2005), Capellari-Jenkins(2002), Capellari-Jenkins(2004), Capellari(1999), Arulampalam-Booth-Taylor(1998), Nicaise-Deblander(2005), Hansen-Wahlberg(2004), Devicienti-Gualteri(2007), Andr  n(2007), Poggi(2007),... who include in a model with unobserved effects employment status and/or household composition and thus assume the latter are strictly exogenous.

⁴⁷ Biewen(2004) and Aasve et al.(2006) are up to now the only ones that estimate simultaneously employment status and household composition in a model of poverty dynamics.

variable is to assume that this is strictly exogenous: Jenkins (2002), Capellari-Jenkins(2004), Meghir-Whitehouse(1997), Arulumpalam-Booth-Taylor(2000), Nicaise-Deblander(2005). Nonetheless, it is true that in the cited models health status is measured at the beginning of the sampling period while in our case it is measured at the end of the sampled period (in 2001). If self-reported health status would not be time-constant, it could be not strictly exogenous and lead to inconsistent estimates. For our dataset, there is no way to test whether it is exogenous or endogenous. Survey information⁴⁸ seems to indicate however that subjective reported health status is rather time-invariant. To avoid discussion, we proxied self-reported health status by life expectancy measured in the first year of the sampling period 1991 by age, education level and gender⁴⁹. This is strictly exogenous and has the advantage of containing more variability than the usual instrument that is the unemployment rate. Finally, after also experimenting with growth rate of GDP, a year dummy for 1992 and combinations of instruments, the best fit and most significant results were obtained with unemployment rate and health expectancy of the head as instruments.

Table 6: multiple-spell discrete-time hazard model with unobserved effects⁵⁰			
Variable	Without heterogeneity Without initial condition	With heterogeneity Without initial condition	With heterogeneity with initial condition
Exit			
Intercept	-17.426** (1.35)	-16.622** (1.40)	-16.307** (1.51)
Duration 1 year	6.278** (0.26)	6.321** (0.27)	5.218** (0.28)
Duration 2 years	5.351** (0.28)	5.451** (0.28)	4.519** (0.28)
Duration 3 years	5.084** (0.28)	5.210** (0.28)	4.378** (0.28)
Duration 4 years	5.117** (0.28)	5.248** (0.28)	4.491** (0.28)
Duration 5 years	4.909** (0.29)	5.055** (0.29)	4.373** (0.29)
Duration 6 years	4.727**	4.879**	4.257**

⁴⁸ Kington-Smith(1998): “Self-reported health status is not used to measure temporary health problems but to include general physical, social and emotional function. Health in old age reflects one's long-term health history. The study's findings show that health status in advanced years is greatly influenced by a history of health that goes back to one's childhood and reaches even beyond personal health status to include the health status of parents and siblings throughout their lives.”

⁴⁹ Deboosere-Gadeyne(2002).

⁵⁰ Since the sample consists of repeated observations on the same household, standard errors are adjusted to account for the dependence at the level of the household.

	(0.29)	(0.30)	(0.30)
Duration 7 years	4.535** (0.31)	4.697** (0.31)	4.128** (0.31)
Duration 8 years	4.395** (0.32)	4.557** (0.32)	4.024** (0.33)
Duration 9 years	4.409** (0.32)	4.659** (0.34)	4.163** (0.35)
AGE	0.399** (0.04)	0.374** (0.04)	0.400** (0.05)
AGE squared	-0.003** (0.0003)	-0.003** (0.0003)	-0.003** (0.0004)
Female head	-0.425** (0.04)	-0.436** (0.04)	-0.487** (0.05)
Member second pillar	0.350** (0.03)	0.376** (0.04)	0.429** (0.04)
Education level head			
Low secondary general	-0.052 (0.06)	-0.053 (0.06)	-0.075 (0.07)
Low secondary technical	0.184* (0.06)	0.213** (0.06)	0.230** (0.07)
Low secondary professional	0.0727 (0.06)	0.077 (0.06)	0.069 (0.07)
High secondary general	0.116 (0.07)	0.135 (0.08)	0.144 (0.09)
High secondary technical	0.207** (0.07)	0.218 (0.08)	0.236 (0.09)
High secondary professional	-0.112 (0.09)	-0.124 (0.10)	-0.160 (0.11)
University	0.298** (0.05)	0.310** (0.05)	0.341** (0.06)
Education level spouse of household head			
Low secondary general	-0.090 (0.06)	-0.124 (0.06)	-0.141 (0.07)
Low secondary technical	-0.052 (0.103)	-0.069 (0.11)	-0.068 (0.12)
Low secondary professional	0.026 (0.06)	0.0165 (0.06)	0.028 (0.07)
High secondary general	0.056 (0.08)	0.065 (0.09)	0.086 (0.10)
High secondary technical	0.073 (0.13)	0.059 (0.14)	0.079 (0.16)
High secondary professional	-0.028 (0.09)	-0.039 (0.10)	-0.034 (0.12)
university	0.219** (0.085)	0.227** (0.09)	0.278** (0.10)
Re-entry			
intercept	-8.882** (1.66)	-8.209** (1.77)	-7.659** (1.88)
Duration 1 year	5.882** (0.33)	5.935** (0.33)	5.450** (0.33)
Duration 2 years	5.183** (0.33)	5.347** (0.33)	4.990** (0.33)
Duration 3 years	4.868** (0.34)	5.115** (0.34)	4.842** (0.34)
Duration 4 years	4.564**	4.873**	4.665**

	(0.34)	(0.34)	(0.34)
Duration 5 years	4.303** (0.34)	4.672** (0.35)	4.522** (0.35)
Duration 6 years	4.213** (0.35)	4.630** (0.35)	4.531** (0.36)
Duration 7 years	4.175** (0.36)	4.634** (0.36)	4.581** (0.37)
Duration 8 years	3.856** (0.38)	4.355** (0.39)	4.348** (0.39)
Duration 9 years	3.945** (0.41)	4.453** (0.41)	4.476** (0.42)
AGE	0.088 (0.05)	0.065 (0.05)	0.061 (0.05)
AGE SQUARED	-0.0009** (0.0004)	-0.0007** (0.0005)	-0.0007 (0.0005)
Female head	0.013 (0.05)	0.039 (0.06)	0.071 (0.07)
Member second pillar	-0.409** (0.05)	-0.478** (0.05)	-0.527** (0.05)
Education household head			
Low secondary general	0.065 (0.08)	0.101 (0.08)	0.124* (0.09)
Low secondary technical	-0.129 (0.07)	-0.167 (0.08)	-0.189 (0.09)
Low secondary professional	-0.0058 (0.07)	-0.026 (0.08)	-0.033 (0.08)
High secondary general	-0.101 (0.08)	-0.142 (0.09)	-0.156 (0.10)
High secondary technical	-0.055 (0.09)	-0.106 (0.10)	-0.137 (0.11)
High secondary professional	0.168 (0.10)	0.183 (0.11)	0.197 (0.12)
university	-0.313** (0.07)	-0.351** (0.07)	-0.381** (0.08)
Education level spouse			
Low secondary general	0.221** (0.07)	0.280** (0.08)	0.316** (0.09)
Low secondary technical	-0.045 (0.12)	-0.041 (0.041)	-0.033 (0.15)
Low secondary professional	-0.024 (0.08)	-0.006 (0.08)	-0.002 (0.09)
High secondary general	0.046 (0.11)	0.080 (0.123)	0.102 (0.13)
High secondary technical	-0.025 (0.15)	0.005 (0.16)	0.0039 (0.18)
High secondary professional	0.142 (0.11)	0.192 (0.12)	0.210 (0.14)
university	-0.358 (0.10)	-0.366 (0.11)	-0.388 (0.12)
Initial condition ⁵¹			
Life expectancy head			-5.199**

⁵¹ The initial condition equation also includes all the explanatory variables in the poverty transition equations but these are not displayed to save space.

			(1.33)
Unemployment rate			21.605** (7.25)
Covariance parameter estimates of unobserved effects			
		$\rho_1 = -0.59$ (0.04)	$\rho_1 = -0.95$ (0.05) $\rho_2 = 0.105$ (0.04) $\rho_3 = -0.035$ (0.04)
Formal likelihood ratio tests of significance of covariance parameter estimates			
		$H_0: \rho_1 = 0$ $\chi^2_1 = 9987; \text{Pr} > \chi^2_1 < 0.0001$	$H_0: \rho_1 = 0$ $\chi^2_1 = 293.89; \text{Pr} > \chi^2_1 < 0.0001$ $H_0: \rho_2 = 0$ $\chi^2_1 = 3.45; \text{Pr} > \chi^2_1 = 0.05$ $H_0: \rho_3 = 0$ $\chi^2_1 = 0.40; \text{Pr} > \chi^2_1 = 0.52$ $H_0: \rho_1 = \rho_2 = 0$ $\chi^2_2 = 366.33; \text{Pr} > \chi^2_1 < 0.0001$ $H_0: \rho_2 = \rho_3 = 0$ $\chi^2_2 = 3.45; \text{Pr} > \chi^2_1 = 0.17$ $H_0: \rho_1 = \rho_2 = \rho_3 = 0$ $\chi^2_3 = 400.25; \text{Pr} > \chi^2_1 < 0.0001$
-2LogLikelihood	461013	459184	405304
Number observations	72844	72844	72844
* denotes significance at 5% level, ** denotes significance at 1% level; standard errors in parenthesis .The reference person is a male head with primary school education, no member of second pillar, living in Antwerp, with spouse with primary school education in 1992. Year and province dummies are included in all equations.			

We primarily compare the coefficients of duration dummies across the three scenarios. When taking into account unobserved heterogeneity, the coefficients become larger. However the duration dummies remain strongly significant and decrease with duration suggesting a genuine causal effect of duration dependence. The unobserved effects do not really change the coefficients of the individual characteristics what means that the unobserved effects are not correlated with the already included individual characteristics. As is typically the case, taking into account the sample selection bias, when going from the second to the third column, leads to a reduction of the estimated coefficients: the exogeneity hypothesis leads to overestimate both size and significance of the estimated coefficients.

Most of the individual characteristics that are significant, like education level, female head and membership second pillar, have opposite signs in the exit and re-entry equations. As in VanKerm(2004), Capellari-Jenkins(2002) and Stewart-Swaffield(1999), the unobserved effects that lead to poverty exit are also decreasing poverty re-entry since ρ_1 has a negative sign ($\rho_1=-0.59$) and is strongly significant. This means that besides the observable characteristics that reduce exit and increase reentry there are also in addition unobserved effects that lead to the same kind of persistence in poverty. This remains so if we take in addition the initial condition into account ($\rho_1=-0.95$).

The positive sign of ρ_2 indicates that the unobserved effects that make that individuals are likely to be initially poor are also increasing the risk of poverty exit. This sign is interpreted by Stewart-Swaffield(1999), Jenkins-Capellari(2002) and Capellari(1999) as follows: given that it measures the correlation between the probability of having a poverty transition and being initially poor, the negative sign is analogous to a negative coefficient in the regression of poverty transitions on poverty status, i.e. Galtonian regression towards the mean. Finally ρ_3 has a negative sign meaning that the unobserved effect that makes that individuals are likely to be initially poor are also decreasing the risk of poverty reentry but is not significant.

A formal likelihood ratio test of significance of the covariance parameter estimates confirmed that the model that allows for correlation between the unobserved effects is clearly to be preferred to the one that does not, what we interpret as evidence of unobserved heterogeneity and an initial condition problem. It also shows that the hypothesis that only $\rho_3=0$ cannot be rejected.

Up to now, we assumed the population is a group of homogenous individuals and interpreted duration dependence as related to work disincentives, stigma or depreciation of human capital. One might argue these arguments concern primarily those covered by the social security system of the employed while the mechanism behind the observed poverty among the self-employed may in addition be related to the non-declaration of

incomes. To check whether the latter mechanism would dominate our results, we estimated the model separately for those who have been employed and those who have been self-employed. Negative dependence in poverty remains a true phenomenon.

5. Concluding remarks

The matching of the National Register with the Income Tax Returns and Census provided evidence of strong income mobility: 1) Every year about 14% of the Belgian civil population is out of the Income Tax Returns, while only 4.9% of households do not appear for any year 1991-2002 in the Income Tax Returns; 2) 37% of the Belgian elderly experience poverty once over a period of 12 years what is much larger than the 12% of Belgian elderly that is poor in a given year.

About 30% of those who become poor leave poverty already after one year and are only transitory poor. The bulk of the elderly poor are however persistently poor. The question arises whether this persistence in poverty is true or spurious. The estimation of a multiple spell discrete-time hazard model, controlling for unobserved effects and a significant initial condition problem, showed a genuine causal effect of duration dependence. One does not know a lot of the mechanism that lies behind genuine persistence in poverty. It has been suggested that persistence may be due to depreciation of human capital or adverse work incentives. The latter illustrates the poverty trap: people may be given a financial incentive not to work while at the same time they slip into poverty. This suggestion sounds reasonable since in Belgium elderly unemployed are exempted from the search for a job and thus easily exposed to depreciation of human capital and employers are reluctant to invest in the human capital of elderly workers. In addition, in Belgium both employers and the government design retirement pathways that give elderly strong incentives to leave the labour market as soon as possible.

6. Bibliography

Aasve, A. - Burgess, S. - Dickson, M.- Propper, C.(2006), “Modelling poverty by not modelling poverty: an application of a simultaneous hazards approach to the UK”, Working Paper 2005-26, ISER, University of Essex.

Allison, P.D. (1984), *Event history analysis. Regression for longitudinal event data*, Sage University Paper Series (Quantitative applications in social sciences).

Andren, T.(2007), “The persistence of welfare participation”, Discussion Paper 3100, IZA, Bonn.

Antolin, P.- Dang, T.- Oxley, H.(1999), « Poverty dynamics in four OECD countries », Economics Department Working Paper 212, OECD, Paris.

Arranz, J.M. – Cantto, O.(2006), “The dynamics of poverty in Spain: the relevance of considering multiple-spells and their accumulation”, Discussion Paper, University of Alcala.

Arulampalam, W. – Booth, A. – Taylor, M.(2000), “Unemployment persistence”, *Oxford Economic papers* 52(1), 24-50.

Bane, M.J. – Ellwood, D.T.(1986), “Slipping into and out of poverty: the dynamics of spells”, *Journal of Human Resources* 21(1), 1-23.

Bardasi, E. – Jenkins, S. – Rigg, J.(2000), “Retirement and the economic well-being of the elderly: a British perspective”, ISER, University of Essex.

Biewen, M.(2003), “Who are the chronic poor? Evidence on the extent and the composition of chronic poverty in Germany”, Discussion Paper 350, DIW, Berlin.

Biewen, M.(2004), “Measuring state dependence in individual poverty status: are there feedback effects to employment decisions and household composition?”, Discussion paper 1138, IZA, Bonn.

Blöndal, S. - Scarpetta, S.(1998), “The retirement decision in OECD countries”, Economics Department Working Paper 202, OECD, Paris.

Boldrin, M. – Dolado, J. J.- Jimeno, J.F.- Perrachi F.(1999), “The Future of Pension in Europe”, *Economic Policy* 29, 289-320.

Bourreau-Dubois, C. – Jeandidier, B. – Berger, F.(2003), “Poverty dynamics, family events, labour market events in Europe: are there any differences between women and men?”, Conference of the European Panel Users Network, ISER, University of Essex.

Callens, M. – Croux, C.- Avramov, D.(2003), “Poverty dynamics in Europe: a multi-level discrete-time recurrent hazard analysis”, EPUNET 2005 Conference, University of Essex.

Cantillon, B. - De Lathouwer, L.- Marx, I.- Van Dam, R.- Van den Bosch, K.(1999), “Sociale indicatoren 1976-1997”, CSB bericht, University of Antwerp.

Capellari, L.(1999), "Low-Wage Mobility in the Italian Labour Market", Warwick Economics Research Paper Series 531, University of Warwick.

Capellari, L. – Jenkins, S.(2002), « Modelling low income transitions », Working Paper 2002-8, ISER, University of Essex.

Capellari, L. – Jenkins, S.(2004), “Modelling low pay transition probabilities, accounting for panel attrition, non-response and initial conditions”, Working Paper 2004-08, ISER, University of Essex.

Cappellari, L. – Dorsett, R. – Haile, G.(2007), “State dependence, duration dependence and unobserved heterogeneity in the employment transition over the 50’s”, Working Paper 2007-16, ISER, University of Essex.

Casey, B. – Yamada, A.(2002), “Getting older, getting poorer? A study of the earnings, pensions, assets and living arrangements of older people in nine countries”, Occasional Papers 60, OECD, Paris.

Cockx, B.(1998), « Analysis of transition data by the minimum chi-square method : an application to welfare spells in Belgium », *The Review of Economics and Statistics* 79(3), 392-405.

Deboosere, P. – Gadeyne, S.(2002),”Sterftetafels naar geslacht, gewest en onderwijsniveau in België, 1991-1996”, Working Paper Steunpunt Demografie - Vakgroep Sociaal onderzoek, Free University of Brussels.

Dejemeppe, M. – Cockx, B.(2005), “Duration dependence in the exit rate out of unemployment. Is it true or spurious”, *Journal of Applied Econometrics* 20(1), 1-23.

Deleeck, H. – Cantillon, B. – Meulemans, B. - Van den Bosch, K. (1992), “Some longitudinal results of the belgian socio-economic panel”, *Journal of Income Distribution* 2(2), 211-231.

Devicienti, F.(2002), “Estimating poverty persistence in Britain”, Working Paper 1, Laboratorio R.Revelli, Centre for Employment Studies, Moncalieri.

Devicienti, F.(2001a), “Poverty persistence in Britain: a multivariate analysis using the BHPS, 1991-1997”, *Journal of Economics* 9, 1-34.

Devicienti, F.- Gualtieri, V.(2007), “The dynamics and persistence of poverty: evidence from Italy”, Working Paper 63, Laboratorio R. Revelli, Centre for Employment Studies, Moncalieri.

Dewilde, C.(2004), « Poverty mobility in the belgian and british welfare regimes : the impact of demographic and labour market events », Research Group on Poverty, Social Exclusion and the City, University of Antwerp.

Dorn, D. - Souza, A.(2005), “Early retirement: free choice or forced decision?”, Working paper 1542, CESifo.

Fenge, R. – Pestieau, P. (2005), *Social security and early retirement*, MIT Press, Cambridge.

Fertig, M. – Tamm, M.(2007), ”Does money buy higher scholling? Evidence from secondary school track choice in German”, Working Paper 2007-58, ECINEQ.

Finnie, R. – Sweetman, A.(2002), “Poverty dynamics: empirical evidence for Canada”, School of policy studies, Working paper 29, Queens University.

Fouarge, D. – Layte, R.(2003), “Duration of poverty spells in Europe”, Working Paper 2003-47, EPAG.

Fouarge, D. – Muffels, R.(2003) “The role of European welfare states in explaining resources deprivation”, Working Paper 2003-41, EPAG.

Förster, M.(2000), “Trends and driving factors in income distribution and poverty in the OECD area”, Labour market and Social Policy Occasional papers 42, OECD, Paris.

Grüber, J. – Wise, D.(eds.)(1999), *Social security and retirement around the world* . University of Chicago Press, Chicago.

Grüber, J. – Wise, D.(eds.)(2004), *Social security programs and retirement around the world: Micro-estimation*, University of Chicago Press, Chicago.

Hansen, J. – Wahlberg, R. (2004), “Poverty persistence in Sweden”, Discussion Paper 1209, IZA, Bonn.

Heckman, J.(1981), « The incidental parameters problem and the problem of initial conditions in estimating a discrete-time discrete data stochastic process », in Manski, C.F., McFadden (eds.), *Structural analysis of discrete data with economic applications*, MIT Press, Cambridge.

Heckman, J. – Singer, B.(1984), “A method for minimizing the impact of distributional assumptions in econometric models for duration data”, *Econometrica* 52(2), 271-320.

Jenkins, S.(2000), “Modelling household income dynamics”, *Journal of Population dynamics* 13, 529-567.

Jenkins, S. – Rigg, J.(2001), *The Dynamics of Poverty in Britain*, Department for Work and Pensions Research Report 157, Corporate Document Services, Leeds.

Jenkins, S.(1995), “Easy estimation methods for discrete-time duration models”, *Oxford bulletin of economics and statistics* 57(1), 129-138.

Jousten, A. (2005), “The benefits of separating early retirees from the unemployed: simulation results for Belgian wage earners”, Discussion Paper 5077, CEPR, University of Liege.

Jousten, A. – Lefebvre, M. – Perelman, S.- Pestieau, P.(2008), “The effects of youth unemployment on early retirement », Working Paper 08/30, IMF.

Kington, R. – Smith, J.(1998), “The relationship between socio-economic status and health of elderly”, Research Brief 5020 Labor and Population, RAND.

Lancaster, T.(1990), *The econometric analysis of transition data*, Econometric Society Monographs, Cambridge University Press, Cambridge.

Lillard, L.A. – Willis, R.J.(1978), “Dynamic aspects of earnings mobility”, *Econometrica* 46(5), 985-1012.

Lillard, L.A. – Panis, C.(1998), “Panel Attrition from the Panel Study of Income dynamics: Household Income, Marital Status, and Mortality”, *Journal of Human Resources* 33(2), 437-457.

Lindeboom, M.(1998), “Microeconomic analysis of the retirement decision: the Netherlands”, Economics Department Papers 207, OECD, Paris.

Makovec, M. (2005), “Poverty entry and exit during old-age: comparative evidence from the ECHP”, European Centre for Social Welfare Policy and Research.

McKernan, S.M. - Ratcliffe, C. (2002), “Transition events in the dynamics of poverty”, The Urban Institute, Washington.

Meghir, C. – Whitehouse, E.(1997), “Labour market transitions into retirement”, *Journal of Econometrics* 79, 327-354.

Meyer, B.(1990), “Unemployment insurance and unemployment spells”, *Econometrica* 58(4), 757-782.

Nicaise, I.- DeBlander, R.(2005), “Maatschappelijke keuzen, structurele armoede en sociale kost”, HIVA, University of Leuven.

Nicoletti, C. – Rondinelli, C.(2006), “The (mis)specification of discrete-time duration models with unobserved heterogeneity: a Monte Carlo study”, Working Paper 2006-53, ISER, University of Essex.

Nickell, S.J.(1979), “The effect of unemployment and related benefits on the duration of unemployment”, *Economic Journal* 89(353), 34-49.

OECD(1994), *The OECD Jobs Strategy*, OECD, Paris.

OECD(1997), *Implementing the OECD Jobs strategy: Lessons from Member Countries' experiences*, OECD, Paris.

OECD(2005a), *Extending opportunities: how active social policy can benefit us all*, OECD, Paris.

OECD(2005b), *Increasing financial incentives to work: the role of employment conditional benefits*, Employment Outlook, OECD, Paris.

Otto, B. – Siedler, T.(2003), “Poverty in Western and Eastern Germany – A detailed comparison”, *Economic Bulletin* 2/2003, 71-76.

Perelman, S. – Schleiper, A. – Stevart, M.(1998), « Dix années plus tard, d'un congrès à l'autre: l'apport des statistiques fiscales à l'étude de la distribution des revenus », Treizième Congrès des économistes de langue française, CIfOP, Charleroi.

Poggi, A.(2007), “Does persistence of social exclusion exist in Spain?”, *Journal of Economic Inequality* 5, 53-72.

Stevens, A.H.(1999), “Climbing out of poverty, falling back in: measuring the persistence of poverty over multiple spells”, *Journal of Human Resources* 3, 557-588.

Stewart, M.(2007), “The inter-related dynamics of unemployment and low pay”, *Journal of applied econometrics* 22(3), 511-531.

Stewart, M. – Swaffield, J.K.(1999), “Low pay dynamics and transition probabilities”, *Economica* 66, 23-42.

Sueyoshi, G.(1995), “A class of binary response models for grouped duration data”, *Journal of applied econometrics* 10, 411-431.

Valletta, R.(2005), « Les entrées et les sorties de la pauvreté dans les économies avancées : la dynamique de la pauvreté au Canada, en Allemagne, en Grande-Bretagne et aux Etats-Unis », Série de documents de recherche – Division de la statistique du Revenu, Federal Reserve Bank, Canada.

van den Berg, G.(2001), “Duration models : specification, identification and multiple durations”, in Heckman J., Leamer E.(eds.) *Handbook of Econometrics*, Volume V, North-Holland.

Van Kerm, P.(2004),”Une évaluation économétrique des flux vers et hors de la pauvreté en Belgique”, IRISS Working Paper Series 2004-04, CEPS/INSTEAD.

Wooldridge, J.(2002), *Econometric analysis of cross section and panel data*, MIT Press, Cambridge.

Zaidi, A. – Makovec, M.- Fuchs, M. – Lipszyc, B. –Lelkes, O. – Rummel, M. – Marin, B. – DeVos, K.(2006), “Poverty of elderly people in EU-25”, First Report for the Project “Poverty of elderly people”.

Département des Sciences Économiques
de l'Université catholique de Louvain
Institut de Recherches Économiques et Sociales

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique